APPENDIX A

~~CONSTRAINED SEARCHING OF AN INDEX~~TECHNIQUE FOR RANKING RECORDS OF A DATABASE

APPENDIX B

CROSS-REFERENCE TO RELATED APPLICATIONS

This patent application is a continuation of U.S. Patent Application No. 09/920,851, filed August 3, 2001, now U.S. Patent No. 6,745,194, which is continuation of U.S. Patent Application No. 09/634,465, filed August 7, 2000, now U.S. Patent No. 6,317,741, which is continuation of U.S. Patent Application No. 09/361,383, filed June 26, 1999, now U.S. Patent No. 6,105,019, which is continuation of U.S. Patent Application No. 09/054,439, filed April 3, 1998, now abandoned, which is continuation of U.S. Patent Application No. 08/694,912, filed August 9, 1996, now U.S. Patent No. 5,745,890, all of which are hereby incorporated by reference herein in their entirety.

APPENDIX C

FIELD OF THE INVENTION

~~This invention relates generally to searching an index of a database, and more particularly to searching locations of an index.~~

This invention relates generally to locating records of a database using a search engine, and more particularly to searching an index for records which are more likely to contain information of interest to users.

APPENDIX D

BACKGROUND OF THE INVENTION

~~In the prior art, it has been well known that computer systems can be used index database. The index can subsequently be searched to locate database records.~~

~~In recent years, a unique distributed database has emerged in the form of the World-Wide-Web (Web). The database records of the Web are in the form of pages accessible via the Internet. Here, tens of millions of pages are accessible by anyone having a communications link to the Internet. The pages are dispersed over millions of different computer systems all over the world. Users of the Internet constantly desire to locate specific pages containing information of interest.~~

~~Indices to databases can assume a number of different forms. In order to maximize throughput, the number of procedures which interface with the indices should be kept to a small number. Also, as much work as possible should be performed by low-level highly optimized procedures to minimize switches to slower application level procedures.~~

~~Therefore, it is desired to perform the searching of indices using a small number of interfaces which include search constraints that can resolve locations of information in the same records.~~

In the prior art, it has been well known that computer systems can be used to index databases, and to search the index to locate records qualified by queries. In recent years, a unique distributed database has emerged in the form of the World-Wide-Web (Web). The database records of the Web are in the form of pages accessible via the Internet. Here, tens of millions of pages are accessible by anyone having a communications link to the Internet.

The pages are dispersed over millions of different computer systems all over the world. Users of the Internet constantly desire to locate specific pages containing information of interest. The pages can be expressed in any number of different character sets such as English, French, German, Spanish, Cyrillic, Kanakata, and Mandarin. In addition, the pages can include specialized components, such as embedded "forms," executable programs, JAVA applets, and hypertext.

Moreover, the pages can be constructed using various formatting conventions, for example, ASCII text, Postscript files, html files, and Acrobat files. The pages can include links to multimedia information content other than text, such as audio, graphics, and moving pictures.

Search engines have been provided to allow users to locate Web pages of interest. These search engines typically have a

query interface where the users specify terms and operators which they want to use to qualify pages.

There are a number of problems with locating pages using an index to the Web. First, the number of pages accessible through the Web is very large, so the number of potential qualifying pages is also going to be large. In addition, many Web users are unsophisticated, so in many instances queries are going to be loosely specified, potentially yielding many pages which may not be of interest to the users. The number of qualifying pages many number in the tens of thousands.

It is desired to minimize the number of index entries which need to be searched for query terms which are not likely to yield fruitful results, and maximize the search of the index using query terms that are more likely to locate records of interest to users.

APPENDIX E

SUMMARY OF THE INVENTION

~~The invention provides a computer implemented method for searching of an index of a database in a constrained manner. The information of the database is stored as a plurality of records.~~

~~A unique location is assigned to each indexable portion of information of the database. Index entries are written to a memory so that each index entry includes a word entry representing a unique indexable portion of information, and one or more location entries for each occurrence of the unique indexable portion information.~~

~~The index entries are sorted according to a collating order of the word entries, and sequentially according to the location entries of each index entry.~~

~~A query is parsed to generate a first term and a second term related by an AND logical operator. The AND operator requires that a first index entry corresponding to the first term and a second index entry corresponding to the second term both have locations in the same record to satisfy query.~~

~~The location entries of the first and second index entries are searched subject to one or more constraints which must be satisfied. The constraints are expressed in the general form as $C(a) \leq C(b) + K$, where $C(a)$ means a current location of the~~

~~first index entry, C(b) means a current location of the second index entry, and K is a predetermined constant.~~

~~The constraints are satisfied by reading locations of the second index entry until the current location of the second index entry is at least equal to the current location of the first index entry plus the predetermined constant.~~

A technique for ranking records of a database is disclosed. In one particular exemplary embodiment, the technique may be realized as a method for ranking records of a database located during a search of an index to the database performed in response to a query, wherein the index has a plurality of index entries, and wherein each index entry has a weight. The method comprises scoring each located record according to the number of times portions of information corresponding to query terms occur in each record and the weight of each index entry corresponding to occurring query terms. The method also comprises storing the score and an identifier of each located record in a respective entry of a ranking list. The method further comprises, in response to having searched a portion of the index, determining if any records yet to be located may achieve a score that is higher than the score of any of the records already located and stored in the ranking list based upon query terms corresponding to index entries having a low weight, and, if not, searching the

index using query terms corresponding to index entries having weights higher than the low weight.

In accordance with other aspects of this particular exemplary embodiment, the ranking list may beneficially have a limit on the number of entries stored therein.

In accordance with further aspects of this particular exemplary embodiment, each index entry may beneficially have a word entry corresponding to a unique portion of information of the database. If such is the case, the method may further beneficially comprise assigning the weight to each index entry according to a difference between the number of records indexed and the number of records including the unique portion of information corresponding to the word entry of the index entry.

In accordance with additional aspects of this particular exemplary embodiment, the method may further beneficially comprise ordering the entries of the ranking list according to the scores. If such is the case, the method may further beneficially comprise providing information associated with each located record to a user in the order of the ranking list. The provided information associated with each located record may beneficially be, for example, the score of each located record or the identifier of each located record.

In another particular exemplary embodiment, the technique

may be realized as at least one signal embodied in at least one carrier wave for transmitting a computer program of instructions configured to be readable by at least one processor for instructing the at least one processor to execute a computer process for performing the above-described method.

In yet another particular exemplary embodiment, the technique may be realized as at least one processor readable carrier for storing a computer program of instructions configured to be readable by at least one processor for instructing the at least one processor to execute a computer process for performing the above-described method.

In still another particular exemplary embodiment, the technique may be realized as a system for ranking records of a database located during a search of an index to the database performed in response to a query, wherein the index has a plurality of index entries, and wherein each index entry has a weight. The system comprises at least one processor configured to score each located record according to the number of times portions of information corresponding to query terms occur in each record and the weight of each index entry corresponding to occurring query terms. The system also comprises at least one memory configured to store the score and an identifier of each located record in a respective entry of a ranking list. In

response to having searched a portion of the index, the at least one processor is also configured to determine if any records yet to be located may achieve a score that is higher than the score of any of the records already located and stored in the ranking list based upon query terms corresponding to index entries having a low weight, and, if not, search the index using query terms corresponding to index entries having weights higher than the low weight.

In still another particular exemplary embodiment, the technique may be realized as a system for ranking records of a database located during a search of an index to the database performed in response to a query, wherein the index has a plurality of index entries, and wherein each index entry has a weight. The system comprises means for scoring each located record according to the number of times portions of information corresponding to query terms occur in each record and the weight of each index entry corresponding to occurring query terms. The system also comprises means for storing the score and an identifier of each located record in a respective entry of a ranking list. The system further comprises means for determining, in response to having searched a portion of the index, if any records yet to be located may achieve a score that is higher than the score of any of the records already located

and stored in the ranking list based upon query terms corresponding to index entries having a low weight. The system still further comprises means for searching the index using query terms corresponding to index entries having weights higher than the low weight if no records yet to be located may achieve a score that is higher than the score of any of the records already located and stored in the ranking list based upon query terms corresponding to index entries having the low weight.

In accordance with other aspects of this particular exemplary embodiment, the means for scoring, the means for determining, and/or the means for searching may beneficially comprise at least one processor. Also, the means for storing may beneficially comprise at least one memory.

The present disclosure will now be described in more detail with reference to exemplary embodiments thereof as shown in the accompanying drawings. While the present disclosure is described below with reference to exemplary embodiments, it should be understood that the present disclosure is not limited thereto. Those of ordinary skill in the art having access to the teachings herein will recognize additional implementations, modifications, and embodiments, as well as other fields of use, which are within the scope of the present disclosure as described herein, and with respect to which the present

disclosure may be of significant utility.

APPENDIX F

1 - 5 (Cancelled).

6 (New). A method for ranking records of a database located during a search of an index to the database performed in response to a query, the index having a plurality of index entries, each index entry having a weight, the method comprising the steps of:

scoring each located record according to the number of times portions of information corresponding to query terms occur in each record and the weight of each index entry corresponding to occurring query terms;

storing the score and an identifier of each located record in a respective entry of a ranking list; and

in response to having searched a portion of the index, determining if any records yet to be located may achieve a score that is higher than the score of any of the records already located and stored in the ranking list based upon query terms corresponding to index entries having a low weight, and, if not, searching the index using query terms corresponding to index entries having weights higher than the low weight.

7 (New). The method of claim 6, wherein the ranking list has a

limit on the number of entries stored therein.

8 (New). The method of claim 6, wherein each index entry has a word entry corresponding to a unique portion of information of the database, further comprising the step of:

assigning the weight to each index entry according to a difference between the number of records indexed and the number of records including the unique portion of information corresponding to the word entry of the index entry.

9 (New). The method of claim 6, further comprising the step of:

ordering the entries of the ranking list according to the scores.

10 (New). The method of claim 9, further comprising the step of:

providing information associated with each located record to a user in the order of the ranking list.

11 (New). The method of claim 10, wherein the provided information associated with each located record is the score of each located record.

12 (New). The method of claim 10, wherein the provided

information associated with each located record is the identifier of each located record.

13 (New). At least one signal embodied in at least one carrier wave for transmitting a computer program of instructions configured to be readable by at least one processor for instructing the at least one processor to execute a computer process for performing the method as recited in claim 6.

14 (New). At least one processor readable carrier for storing a computer program of instructions configured to be readable by at least one processor for instructing the at least one processor to execute a computer process for performing the method as recited in claim 6.

15 (New). A system for ranking records of a database located during a search of an index to the database performed in response to a query, the index having a plurality of index entries, each index entry having a weight, the system comprising:

at least one processor configured to score each located record according to the number of times portions of information corresponding to query terms occur in each record and the weight

of each index entry corresponding to occurring query terms; and

at least one memory configured to store the score and an identifier of each located record in a respective entry of a ranking list;

wherein, in response to having searched a portion of the index, the at least one processor is also configured to determine if any records yet to be located may achieve a score that is higher than the score of any of the records already located and stored in the ranking list based upon query terms corresponding to index entries having a low weight, and, if not, search the index using query terms corresponding to index entries having weights higher than the low weight.

16 (New). The system of claim 15, wherein the ranking list has a limit on the number of entries stored therein.

17 (New). The system of claim 15, wherein each index entry has a word entry corresponding to a unique portion of information of the database, wherein the at least one processor is also configured to assign the weight to each index entry according to a difference between the number of records indexed and the number of records including the unique portion of information corresponding to the word entry of the index entry.

18 (New). The system of claim 15, wherein the at least one processor is also configured to order the entries of the ranking list according to the scores.

19 (New). The system of claim 18, wherein the at least one processor is also configured to provide information associated with each located record to a user in the order of the ranking list.

20 (New). The system of claim 19, wherein the provided information associated with each located record is the score of each located record.

21 (New). The system of claim 19, wherein the provided information associated with each located record is the identifier of each located record.

22 (New). A system for ranking records of a database located during a search of an index to the database performed in response to a query, the index having a plurality of index entries, each index entry having a weight, the system comprising:

means for scoring each located record according to the number of times portions of information corresponding to query terms occur in each record and the weight of each index entry corresponding to occurring query terms;

means for storing the score and an identifier of each located record in a respective entry of a ranking list;

means for determining, in response to having searched a portion of the index, if any records yet to be located may achieve a score that is higher than the score of any of the records already located and stored in the ranking list based upon query terms corresponding to index entries having a low weight; and

means for searching the index using query terms corresponding to index entries having weights higher than the low weight if no records yet to be located may achieve a score that is higher than the score of any of the records already located and stored in the ranking list based upon query terms corresponding to index entries having the low weight.

23 (New). The system of claim 22, wherein the means for scoring comprises at least one processor.

24 (New). The system of claim 22, wherein the means for storing

comprises at least one memory.

25 (New). The system of claim 22, wherein the means for determining comprises at least one processor.

26 (New). The system of claim 22, wherein the means for searching comprises at least one processor.

APPENDIX G

ABSTRACT OF THE DISCLOSURE

~~A computer implemented method performs constrained searching of an index of a database. The information of the database is stored as a plurality of records. A unique location is assigned to each indexable portion of information of the database. Index entries are written to a memory where each index entry includes a word entry representing a unique indexable portion of information, and one or more location entries for each occurrence of the unique indexable portion information. The index entries are sorted according to a collating order of the word entries, and sequentially according to the location entries of each index entry. A query is parsed to generate a first term and a second term related by an AND logical operator, the AND operator requires that a first index entry corresponding to the first term and a second index entry corresponding to the second term both have locations in the same record to satisfy a query. The location entries of the first and second index entries are searched subject to one or more constraints which must be satisfied. The constraints are expressed as $C(a) \leq C(b) + K$, where $C(a)$ means a current location of the first index entry, $C(b)$ means a current location of the second index entry, and K is a predetermined constant.~~

A technique for ranking records of a database is disclosed. In one particular exemplary embodiment, the technique may be realized as a method for ranking records of a database located during a search of an index to the database performed in response to a query, wherein the index has a plurality of index entries, and wherein each index entry has a weight.